DOCUMENT RESUME

ED 438 327 TM 030 644

AUTHOR	Canivez, Gary L.
TITLE	Stability of Regular Education Teacher Ratings of Normal and Exceptional Students on the Adjustment Scales for Children and Adolescents (ASCA).
PUB DATE	1999-00-00
NOTE	11p ; Paper presented at the Annual Meeting of the National Association of School Psychologists (31st, Las Vegas, NV, April 6-10, 1999).
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Behavior Patterns; *Classification; Disabilities; Elementary Education; *Elementary School Students; Measures (Individuals); *Reliability; *Student Adjustment
IDENTIFIERS	*Adjustment Scales for Children and Adolescents; *Test Retest Reliability

ABSTRACT

The short-term (45-day) stability of the Adjustment Scales for Children and Adolescents (P. McDermott, N. Marston, and D. Stott, 1993) was studied with 51 first and fifth graders, seven of whom were classified as "exceptional/disabled." Significant test-retest reliability coefficients were obtained, and mean differences from test to retest did not exceed 0.5 raw score points. Only two scales (Solitary Aggressive/Impulsive and Lethargic/Hypoactive) showed significant changes across the retest intervals for T scores. Individual variation for some scales was at times quite extreme. Syndromic profile classifications and discriminant classifications were also significantly consistent across the retest interval. (Contains 31 references, 6 tables, and 1 figure.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 438 327

Stability of Regular Education Teacher Ratings of Normal and Exceptional Students on the Adjustment Scales for Children and Adolescents (ASCA)

Gary L. Canivez
Eastern Illinois University

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Gary L. Canivez
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at the 1999
Annual Convention of the National Association of School Psychologists, Las Vegas, NV

BEST COPY AVAILABLE

Stability of Regular Education Teacher Ratings of Normal and Exceptional Students on the *Adjustment Scales for Children and Adolescents* (ASCA)

Gary L. Canivez, *Eastern Illinois University*

Paper presented at the 1999 Annual Convention of the National Association of School Psychologists, Las Vegas, NV

Correspondence regarding this presentation should be addressed to Gary L. Canivez, Ph.D., Department of Psychology, Eastern Illinois University, 600 Lincoln Avenue, Charleston, IL 61920-3099. Dr. Canivez may also be contacted via E-mail at cfglc@eiu.edu or the World Wide Web at <http://www.ux1.eiu.edu/~cfglc>.

Independent investigation of the short term (45 day) stability of the Adjustment Scales for Children and Adolescents is reported. Significant test-retest reliability coefficients were obtained and mean differences from test to retest did not exceed .5 raw score points. Only two scales (Solitary Aggressive-Impulsive & Lethargic/Hypoactive) showed significant changes across the retest intervals for T scores. Individual variation for some scales was at times quite extreme. Syndromic profile classifications and discriminant classifications were also significantly consistent across the retest interval.

School psychologists prefer objective assessment methods which can facilitate a link between assessment and intervention (Reschly & Ysseldyke, 1995). Standardized behavior rating scales and checklists have achieved great popularity among school and clinical psychologists (Merrell, 1994a). Among school psychologists they are the most frequently used instruments in assessing emotional and behavioral difficulties of youths (Stinnett, Havey, & Oehler-Stinnet, 1994). Behavior rating scales are "one of the most efficient, sound, and effective ways ... to identify a referred student's behavioral strengths and weaknesses..." (Knoff, 1995, p. 857). Use of behavior rating scales have also been designated a "best practice" in the assessment of emotional and behavioral disorders (McConaughy & Ritter, 1995).

Behavior rating scales offer, among other advantages, unobtrusive evaluations of students' behavior in the natural social settings such as schools, classrooms, and homes. Within the classroom and other school settings, teachers are natural observers and informants since they have the comparative experience of observing many students across time and varied social contexts. As such, they appear to take a normative perspective in rating difficulties in children. Consequently, teachers have sometimes been considered to be among the most accurate adult raters of child behavior (Kamphaus & Frick, 1996).

The Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993) is a relatively new behavior rating scale designed to assess youth psychopathology based on teacher report of child behaviors in school settings. Evidence of short term (30 school day) stability reported in the ASCA *Manual* was based upon a sample 40, 14-17 year old female students in Pennsylvania (McDermott, 1994). All correlations were significant and there were no significant mean *T* score differences from test to retest.

Previous investigation of the ASCA's short term stability pertained to the obtained *T* scores and their cut score interpretations. In addition to cut score interpretation, McDermott (1994) also presents two *multivariate* methods of interpretation: *Syndromic Profile Classification* and *Discriminant Classification*. Neither of these classification methods have been investigated in relation to their stability.

Syndromic Profile Classification is based on results of the cluster analysis of the standardization sample which produced 22 profile types (McDermott, 1994; McDermott & Weiss, 1993; 1995). *Syndromic Profile Classification* involves comparing a youth's core syndrome *T* scores to the mean *T* scores for one or more of the 22 profile types (14 major types and 8 clinical subtypes) to determine which profile is most similar. Classification of the youth's profile is based on the *generalized distance score* (GDS) method (McDermott, 1994). Canivez (1996, 1998a) and Watkins (1997) automated the calculations for the GDS to assure reliable calculation. Watkins (1997) also provided an additional profile similarity coefficient, $r_{p(k)}$, in his program. The $r_{p(k)}$ coefficient is a special version of Cattell's (1949) $r_{p(k)}$ based on a formula provided by Tatsuoka (1974).

Discriminant Classification is based on results of discriminant function analysis conducted in which the ASCA was found to correctly classify normal from socially/emotionally disturbed youths (McDermott, 1994; McDermott, Watkins, Sichel, Weber, Keenan, Holland, & Leigh, 1995). *Discriminant Classification* involves applying the youth's six core syndrome *T* scores to the two linear discriminant function regression equations to determine which group the youth in question is most likely to belong. The youth's profile is classified as most similar to the group (normal vs. socially/emotionally disturbed) which results in the *higher* discriminant score. Canivez (1996, 1998a) and Watkins (1997) also provide automated calculation of *Discriminant Classifications*.

Given the potential diagnostic applications of the ASCA, independent assessment of the ASCA's stability and a more diverse sample is needed. Consequently, the purpose of this study was to examine the short term test-retest stability of the ASCA with a more diverse sample of students enrolled in regular educational programs. Additionally, the present study sought to replicate and extend previous results by examining the stability of syndromic profile classifications and discriminant classifications in addition to stability of ASCA raw scores and *T* scores. Stability data for syndrome profiles and discriminant classifications have yet to be reported.

Method

Participants

The sample included 51 (27 male, 24 female) students attending a public elementary school in a suburban location of a large southwestern metropolitan area. Students attended first (n = 26) or fifth (n = 25) grade classes and were primarily Caucasian (94%). Normal (n = 44) and exceptional/disabled (n = 7) students were twice rated on the ASCA by their regular education classroom teachers. Disabled students were previously classified by multidisciplinary evaluation teams as prescribed by federal and state regulations and were attending their regular education classroom the majority of the school day.

Instrument

The Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993) is an objective behavior rating instrument completed by a student's classroom teacher and designed for use with all noninstitutionalized youths ages 5-17 (grades K-12). The ASCA consists of 156 behavioral descriptions within 29 specific situations where teachers may observe student's behaviors. Of the 156 items, 96 are scorable for dimensions of psychopathology and based on factor analyses, singularly assigned to one of six core syndromes (Attention-Deficit/Hyperactive [ADH], Solitary Aggressive-Provocative [SAP], Solitary Aggressive-Impulsive [SAI], Oppositional Defiant [OPD], Diffident [DIF], and Avoidant [AVO]) or two supplementary syndromes (Delinquent [DEL] and Lethargic/Hypoactive [LEH]). The six core syndromes are combined to form two composite indexes: Overactivity (OVR: Attention-Deficit Hyperactive, Solitary Aggressive-Provocative, Solitary Aggressive-Impulsive, and Oppositional Defiant syndromes) and Underactivity (UNR: Diffident and Avoidant syndromes). Raw scores are converted to normalized T scores (by area conversion) based on the nationally representative standardization sample. ASCA was normed on a random, representative national sample of 1,400 youths, blocked according to gender, age, and grade level and stratified proportionately according to national region, community size, race/ethnicity, parent education, family structure, and handicapping condition.

Extensive reliability and validity evidence is provided in the ASCA *Manual* (McDermott, 1994). Internal consistency estimates for the total standardization sample ranged from .68 to .86 for the six core syndromes and two supplementary syndromes. Alpha coefficients equalled .92 for the Overactivity scale and .82 for the Underactivity scale. Test-retest reliabilities over a 30 school day interval ranged from .66 to .91 for the six core syndromes (n = 40). Test-retest correlations equalled .75 for the Overactivity scale and .79 for the Underactivity scale. Studies of interobserver agreement for the core syndromes and global adjustment scales found significant agreement in both level and pattern (McDermott, 1994; Watkins & Canivez, 1997). Exploratory and confirmatory analyses support the factor structure at the item, core syndrome, and second-order levels. Convergent and divergent validity studies comparing the ASCA with the Conners Teacher Rating Scale (CTRS; Trites, Blouin, & Laprade, 1982) and the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) found significant correlations among similar psychological dimensions (McDermott, 1994). In general, psychometric characteristics of the ASCA are acceptable and meet standards for both group and individual decision making (Salvia & Ysseldyke, 1995).

Procedure

Two regular eduction classroom teachers in a suburban southwest metropolitan public school district volunteered to participate in the present study. The teachers rated individual students attending their class using the ASCA following the standard administration procedures. All ASCA rating forms were returned to this author who scored them according to standard procedures (McDermott, 1994). The teachers again rated the same students 45 days later. Core Syndrome, Supplementary Syndrome, and global Adjustment Scale T scores were obtained from the ASCA *Manual*.

Syndromic Profile Classifications were made using the generalized distance score (GDS) method according to the ASCA *Manual* (Canivez, 1996, 1998a; McDermott, 1994) and the $r_{p(k)}$ method (Watkins, 1997). The GDS is a measure of profile similarity (dissimilarity) by examining deviations of a youth's core syndrome T scores from the average T scores for a specified group (ASCA profile type). The youth's profile is classified as most similar to the ASCA profile type which results in the *smallest* GDS. With the $r_{p(k)}$ method, the youth's profile is classified as most similar to the ASCA profile type which results in the *highest* $r_{p(k)}$ value.

Broad classifications based on syndromic profile types were also made and examined for stability. As indicated in the ASCA *Manual* (pp. 24-25), profile Type 1 is classified *Adjusted*, Types 2 through 5 are classified *Adequately Adjusted*, Types 6 through 12 are classified *Marginally Adjusted*, Types 13 through 18 are classified *At Risk*, and Types 19 through 22 are classified *Maladjusted*. Stability among these five classification categories was assessed. Additional reductions of these five categories were performed to investigate further effects on stability. *Adjusted, Adequately Adjusted, and Marginally Adjusted* groups were pooled into an *Adjusted* category and stability compared with the *At-Risk* and *Maladjusted* groups. Finally, the *At-Risk* and *Maladjusted* groups were pooled into a classification termed *Not Adjusted* and stability for *Adjusted* and *Not Adjusted* groups were examined. Table 1 presents the classifications of each of the 22 syndromic profiles into the 5, 3, and 2 category groupings.

Discriminant Classifications were also made according to the ASCA *Manual* using linear discriminant classification equations (Canivez, 1996, 1998a; McDermott, 1994, p. 29). Profiles were classified normal or socially/emotionally disturbed based on the equation resulting in the highest discriminant score. In the case of tied results, the profile was classified as normal. Stability for discriminant classifications was also investigated.

Data Analyses

Pearson product-moment correlation coefficients between first and second ratings were calculated for raw scores and T scores obtained for the ASCA Core Syndromes, Supplementary Syndromes, and overall Adjustment Scales. Dependent t-tests were conducted to investigate changes in ratings from test to retest. Effect strengths of rating changes across the retest interval were estimated using η^2, an index of the proportion of variability explained by the effect across the retest interval (Kiess, 1996). Individual variation in scores across the test-retest interval was explored through frequency distributions for both raw and T scores.

Table 1
Broad classifications of the 22 Syndromic Profiles into 5, 3, and 2 classification levels

Syndromic Profile Type	5 Level Classification	3 Level Classification	2 Level Classification
1. Good Adjustment	Adjusted	Adjusted	Adjusted
2. Adequate Adjustment w/Inhibition	Adequate Adjustment	Adjusted	Adjusted
3. Adequate Adjustment w/Disruptiveness	Adequate Adjustment	Adjusted	Adjusted
4. Adequate Adjustment w/Apprehension	Adequate Adjustment	Adjusted	Adjusted
5. Adequate Adjustment w/Indifference	Adequate Adjustment	Adjusted	Adjusted
6. Marginal Adjustment w/Withdrawal	Marginal Adjustment	Adjusted	Adjusted
7. Marginal Adjustment w/Motivation Deficit	Marginal Adjustment	Adjusted	Adjusted
8. Marginal Adjustment w/Avoidance	Marginal Adjustment	Adjusted	Adjusted
9. Marginal Adjustment w/Attention Seeking	Marginal Adjustment	Adjusted	Adjusted
10. Marginal Adjustment w/Moodiness	Marginal Adjustment	Adjusted	Adjusted
11. Marginal Adjustment w/Nonparticipation	Marginal Adjustment	Adjusted	Adjusted
12. Marginal Adjustment w/Dependency	Marginal Adjustment	Adjusted	Adjusted
13. Undersocialized Aggressive	At-Risk	At-Risk	Not Adjusted
14. Oppositional	At-Risk	At-Risk	Not Adjusted
15. Provocative, Attention Seeking	At-Risk	At-Risk	Not Adjusted
16. Provocative, Manipulative	At-Risk	At-Risk	Not Adjusted
17. Impulsive Aggressive	At-Risk	At-Risk	Not Adjusted
18. Attention-Deficit Hyperactive	At-Risk	At-Risk	Not Adjusted
19. Instrumental Aggressive	Maladjusted	Maladjusted	Not Adjusted
20. Defiant Aggressive	Maladjusted	Maladjusted	Not Adjusted
21. Avoidant	Maladjusted	Maladjusted	Not Adjusted
22. Schizoid with Depressed Mood	Maladjusted	Maladjusted	Not Adjusted

Unlike the Core Syndromes, Supplementary Syndromes, and overall Adjustment Scale *T* scores; *Syndromic Profile Classifications* and *Discriminant Classifications* are nominal scale variables. When investigating stability or agreement on nominal scale or categorical variables, kappa (κ) is an appropriate statistic (Cohen, 1960; Fleiss, 1981, McDermott, 1988). Kappa provides an index of agreement beyond chance agreement and is interpreted much like a correlation coefficient. Kappa coefficients and statistical tests of kappa were calculated using templates created for the ASCA (Canivez, 1998b) to estimate the stability of the Syndromic Profile Classifications (22 specific Profile Types and 5, 3, and 2 broad classifications) and Discriminant Classifications.

Results

Raw score and *T* score test-retest correlations, descriptive statistics, dependent *t*-tests, and retest interval effect strengths (η^2) for Core Syndromes, Supplementary Syndromes, and global Adjustment Scales are presented in Table 2 and Table 3, respectively. All test-retest reliability coefficients (except LEH raw scores) were significant. Test-retest reliability coefficients ranged from .23 to .82 (*Mdn* = .72) for raw scores and ranged from .38 to .79 (*Mdn* = .66) for *T* scores. Most syndrome raw scores and *T* scores showed no significant mean changes across the retest interval. For raw scores, the SAP, SAI, DIF, and LEH syndromes and the UNR adjustment scale showed significant changes across the retest interval. Effect strengths were small to moderate. For the *T* scores, only the SAI and LEH syndromes showed significant changes across the retest interval. The effect strength for SAI was small however, the effect strength for LEH was moderate.

Figure 1 presents the mean ASCA profiles from the first and second testing (ratings). As seen in Figure 1, mean ASCA profiles are quite similar in their pattern across the retest interval. Significant changes (decreases) in *T* scores were only observed for the SA(I) and LEH syndromes. These comparisons relate to nomothetic stability but do not address the idiographic perspective.

Individual variations in raw scores and *T* scores across the test-retest interval are presented in Tables 4 and 5, respectively, and focus on an idiographic perspective of syndrome stability. As illustrated in Table 5, the majority of individuals showed changes in raw scores ±2 points or less. Changes also tended to be skewed in the direction of more students showing decreases in raw scores across the retest interval. For the Adjustment Scales, 73.6% of students showed changes ±2 raw score points or less on the OVR scale while 83% showed changes ±2 points or less on the UNR scale. For the Core Syndromes and Supplementary Syndromes, the following percentages of students showing ±2 raw score points or less were observed: ADH (79.2%), SA(P) (98.1%), SA(I) (100%), OPD (96.3%), DIF (94.4%), AVO (90.5), DEL (98.1), and LEH (96.2%).

Table 5 illustrates the idiographic comparison of ASCA syndrome stability as it related to the *T* scores. Although there was a general tendency for students to remain relatively constant in their *T* scores, some individuals showed fairly large increases or decreases in their *T* scores for various syndromes. In comparing these changes across time to the standard errors of measurement presented in the ASCA *Manual* (McDermott, 1994) and Table 5, 66.6% of students fell within the 95% confidence interval for the OVR scale while 76.5% of students fell within the 95% confidence interval on the UNR scale. For the Core Syndromes and

Table 2
Test-retest Correlation Coefficients, Descriptive Statistics, t-tests, and Effect Strengths for ASCA Raw Scores

			First Testing			Second Testing				
Syndrome/Scale	*r*	*p*	*M*	*SD*	*Mdn*	*M*	*SD*	*Mdn*	*t*	η^2
Core Syndromes										
ADH	.66	.001	2.06	2.86	1	1.73	2.02	1	1.11	.02
SA(P)	.82	.001	.33	.74	0	.59	1.15	0	2.64**	.12
SA(I)	.43	.002	.10	.30	0	.02	.14	0	2.06*	.08
OPD	.69	.001	1.00	1.77	0	.76	1.32	0	1.30	.03
DIF	.80	.001	1.08	1.74	0	.88	1.32	0	1.32	.03
AVO	.37	.007	1.06	1.52	1	.78	1.08	0	1.31	.03
Supplemental Syndromes										
DEL	.75	.001	.52	1.12	0	.15	.46	0	2.29*	.17
LEH	.23	.104	.43	.73	0	.08	.27	0	3.52***	.20
Adjustment Scales										
OVR	.81	.001	3.49	5.17	1	3.10	4.06	1	.92	.02
UNR	.77	.001	2.14	2.53	1	1.67	1.80	1	2.06*	.08

Note. ADH = Attention Deficit-Hyperactive, SA(P) = Solitary Aggressive (Provocative), SA(I) = Solitary Aggressive (Impulsive), OPD = Oppositional Defiant, DIF = Diffident, AVO = Avoidant, DEL = Delinquent, LEH = Lethargic (Hypoactive), OVR = Overactivity, UNR = Underactivity.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3
Test-retest Correlation Coefficients, Descriptive Statistics, t-tests, and Effect Strengths for ASCA T Scores

			First Testing			Second Testing				
Syndrome/Scale	*r*	*p*	*M*	*SD*	*Mdn*	*M*	*SD*	*Mdn*	*t*	η^2
Core Syndromes										
ADH	.59	.001	50.27	9.18	52	49.31	8.53	52	.85	.01
SA(P)	.79	.001	49.86	9.37	45	50.73	9.95	45	.97	.02
SA(I)	.43	.002	49.16	6.61	47	47.43	3.08	47	2.06*	.08
OPD	.75	.001	51.98	11.86	43	49.86	9.62	43	1.93	.07
DIF	.63	.001	47.92	9.76	40	47.53	9.09	40	.34	.00
AVO	.44	.001	51.78	9.99	57	49.75	9.22	42	1.43	.04
Supplemental Syndromes										
DEL	.75	.001	50.30	11.33	45	47.78	8.02	45	1.75	.11
LEH	.38	.007	50.33	9.15	44	45.41	4.89	44	4.09**	.25
Adjustment Scales										
OVR	.69	.001	50.88	9.11	50	50.00	8.76	50	.90	.02
UNR	.74	.001	49.84	9.77	51	48.45	8.88	51	1.45	.04

Note. ADH = Attention Deficit-Hyperactive, SA(P) = Solitary Aggressive (Provocative), SA(I) = Solitary Aggressive (Impulsive), OPD = Oppositional Defiant, DIF = Diffident, AVO = Avoidant, DEL = Delinquent, LEH = Lethargic (Hypoactive), OVR = Overactivity, UNR = Underactivity.
*$p < .05$. **$p < .001$.

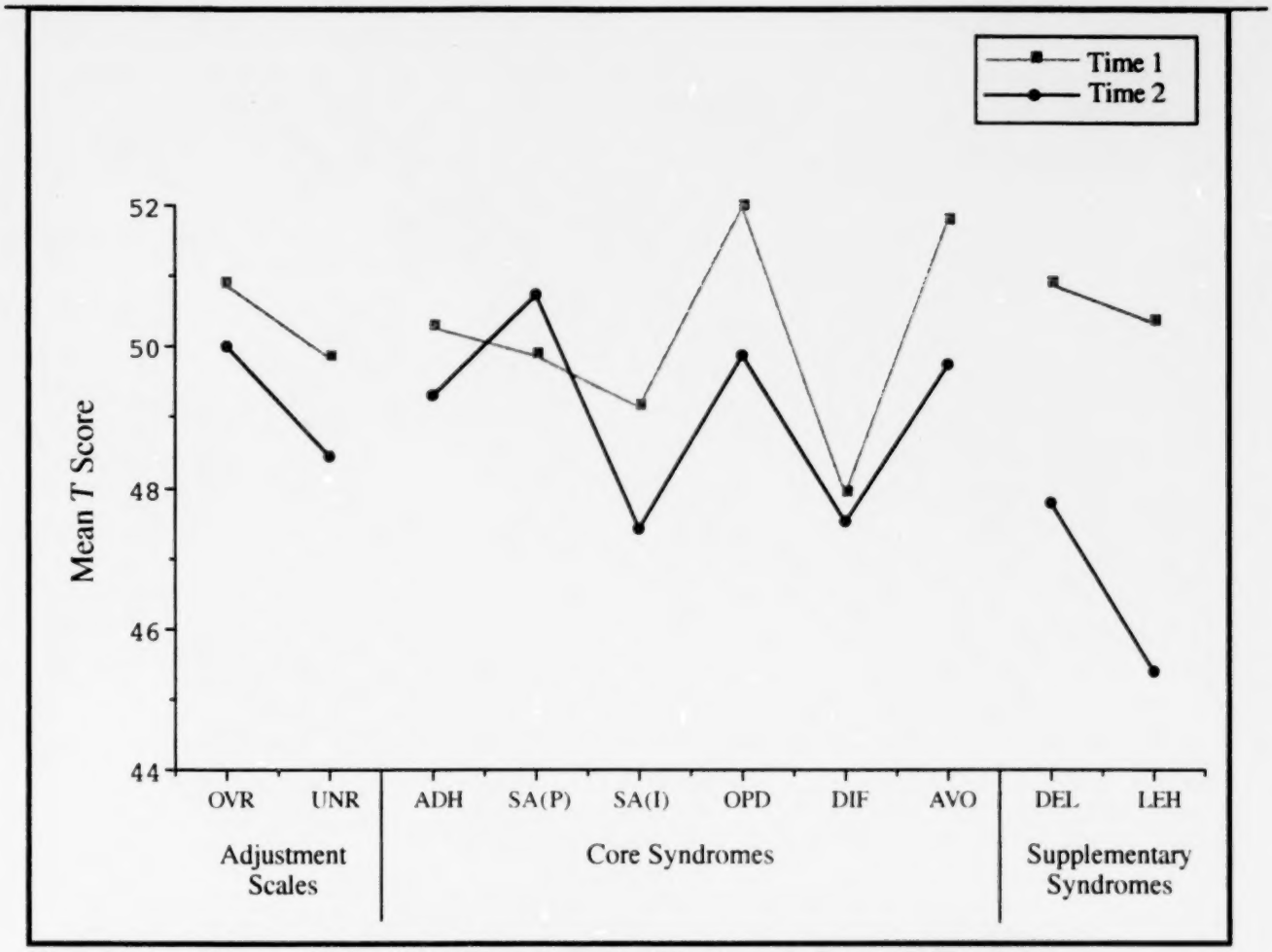

Figure 1. Mean ASCA Profiles for the First and Second Testing.

Supplementary Syndromes, the following percentages of students fell within the 95% confidence interval: ADH (72.7%), SA(P) (88.9%), SA(I) (92.2%), OPD (86.4%), DIF (72.8%), AVO (68.7%), DEL (92.6%), and LEH (74.5%).

Stability of the 22 Syndromic Profile Classifications and their resulting 5, 3, and 2 Broad Classifications are summarized in Table 6. The GDS method and the r_{pik} method produced virtually identical results; however, the GDS method produced consistently greater stability in each level of analysis. As expected, the fewer classifications made, the greater the agreement observed. All kappa coefficients were significant, indicating that classifications of profiles from Time 1 to Time 2 were stable. For the 22 Syndromic Profiles, agreement was fair to moderate while agreement for 5 Broad Classifications were moderate to substantial (Everitt & Hay, 1992; Landis & Koch, 1977). Agreement for 3 and 2 Broad Classifications were almost perfect (Everitt & Hay, 1992; Landis & Koch, 1977).

The third and final stability investigation involved the Discriminant Classification (based on the linear method) made at Time 1 and Time 2. The nominal scale agreement statistics for Discriminant Classifications are also presented in Table 6. As with the Syndromic Profile Classifications, the Discriminant Classifications showed significant and almost perfect agreement (Everitt & Hay, 1992; Landis & Koch, 1977) Time 1 to Time 2 (κ = .56, Z = 4.19, $p < .00003$). Of the 51 students rated at Time 1 and Time 2, 40 (78%) were classified "Normal" at Time 1 and Time 2 while 5 (10%) were classified "SED" at Time 1 and Time 2. Five students (10%) were classified "SED" at Time 1 and "Normal" at Time 2 while 1 (2%) student was classified as "Normal" at Time 1 but "SED" at Time 2.

Discussion

The present study investigated the short term stability of the Adjustment Scales for Children and Adolescents with a sample of students attending regular education classrooms and rated by their regular education classroom teacher. Test-retest correlation coefficients across the 45 day interval were significant but lower in magnitude than those found in a comparable study reported in the ASCA *Manual* (McDermott, 1994). The present study found significant T score changes across the retest interval for only the SA(I) and LEH syndromes which showed small to moderate effect sizes, whereas McDermott (1994) reported no significant changes across the retest interval for any of the syndromes.

Table 4
Frequency Distributions (Percent) of ASCA Raw Score Changes at Second Testing

Δ	OVR	UNR	ADH	SA(P)	SA(I)	OPD	DIF	AVO	DEL	LEH
-11	1 (1.9)									
-10	-									
-9	-									
-8	-									
-7	2 (3.8)					1 (1.9)				
-6	-		2 (3.8)			-		1 (1.9)		
-5	-	2 (3.8)	-			-		1 (1.9)		
-4	3 (5.7)	-	3 (5.7)			-		-		
-3	2 (3.8)	5 (9.4)	3 (5.7)			1 (1.9)	2 (3.8)	2 (3.8)	1 (1.9)	2 (3.8)
-2	5 (9.4)	5 (9.4)	3 (5.7)			2 (3.8)	3 (5.7)	4 (7.5)	4 (7.5)	1 (1.9)
-1	7 (13.2)	8 (15.1)	7 (13.2)	1 (1.9)	4 (7.5)	6 (11.3)	9 (17.0)	6 (11.3)	1 (1.9)	10 (18.9)
0	18 (34.0)	22 (41.5)	21 (39.6)	43 (81.1)	49 (92.5)	38 (71.7)	31 (58.5)	29 (54.7)	48 (90.6)	40 (75.5)
1	8 (15.1)	8 (15.1)	7 (13.2)	5 (9.4)		2 (3.8)	6 (11.3)	7 (13.2)		
2	1 (1.9)	1 (1.9)	4 (7.5)	3 (5.7)		3 (5.7)	1 (1.9)	2 (3.8)		
3	2 (3.8)	2 (3.8)	1 (1.9)	1 (1.9)			1 (1.9)	1 (1.9)		
4	2 (3.8)		1 (1.9)							
5	1 (1.9)		1 (1.9)							
6	-									
7	-									
8	-									
9	1 (1.9)									

Note. Δ = Raw score change from first to second ratings, OVR = Overactivity, UNR = Underactivity, ADH = Attention Deficit-Hyperactive, SA(P) = Solitary Aggressive (Provocative), SA(I) = Solitary Aggressive (Impulsive), OPD = Oppositional Defiant, DIF = Diffident, AVO = Avoidant, DEL = Delinquent, LEH = Lethargic (Hypoactive). Percents presented in parentheses.

The test-retest reliability coefficients in the present study, although significant, are generally lower than those found for other teacher report child behavior rating scales across a similar retest interval (Achenbach, 1991; Merrell, 1994b; Naglieri, LeBuffe, & Pfeiffer, 1993; Reynolds & Kamphaus, 1992). One possible reason for these differences may be in how the items are scored. The ASCA items are dichotomously (0-Absent, 1-Present) scored while other behavior rating scales like the *Child Behavior Checklist 91:Teachers Report Form* (CBCL-TRF; Achenbach, 1991), the *Behavior Assessment System for Children: Teacher Rating Scales* (BASC-TRS; Reynolds & Kamphaus, 1992), the *Preschool and Kindergarten Behavior Scales* (PKBS; Merrell, 1994b), and the *Devereux Behavior Rating Scale-School Form* (Naglieri, LeBuffe, & Pfeiffer, 1993) have items which are scored on a 4 or 5 point continuum, thus increasing variability at the item level as well as in the total scale or syndrome. Given this situation, one would expect higher correlations as a function of the greater available item variability.

This is the first study attempting to investigate the stability of the two multivariate interpretive classification methods presented in the ASCA *Manual*. Results found that the 22 syndromic profile classifications and their resulting 5, 3, and 2 level broad classifications all demonstrated significant agreement across the retest interval indicating significant temporal stability. This is an encouraging and important finding to the extent that one would expect that the profile generated in a behavioral or psychopathology measure should be relatively stable over the short term retest interval investigated in this study. This study also presents a method to investigate such agreement in classification over time for nominal scale classifications through the use of kappa and its significance test (Cohen, 1960; Fleiss, 1981). The same method would also be appropriate for examining classification agreements between two independent raters (interrater reliability) of the ASCA.

Caution should be exercised in interpreting the results of this study as it is based on a small sample of students whom are not representative of the population at large nor were they randomly selected. Generalizability of these results is certainly limited as the sample was predominantly Caucasian (94%) and included students in only the first and fifth grade. Additionally, only two teachers provided ratings of their students and these two teachers do not adequately represent the population of teachers whom might complete the ASCA. Future studies should continue to investigate the temporal stability of the ASCA in a similar manner as this study and incorporate more diverse and representative student and teacher samples and investigate longer test-retest time intervals. Future studies should also continue to investigate reliability of syndromic profile classifications and discriminant classifications.

Table 5

Frequency Distributions (Percent) of ASCA T Score Test-Retest Changes

Δ	OVR SE_m= 2.8	UNR 4.4	ADH 3.9	SA(P) 4.4	SA(I) 3.7	OPD 4.8	DIF 4.7	AVO 5.5	DEL 5.0	LEH 4.8
-36						1 (2.0)				
-35						-				
-34						-				
-33						-				
-32						-				
-31						-				
-30						-				
-29						-		1 (2.0)	2 (3.7)	
-28						-		-	-	
-27						-		-	2 (3.7)	
-26						-		-	-	
-25						-		-	-	2 (3.9)
-24						-		-	-	-
-23			1 (2.0)			-		-	-	-
-22			-	1 (2.0)	4 (7.8)	-		1 (2.0)	-	-
-21			-	-	-	-		-	-	1 (2.0)
-20			-	-	-	1 (2.0)		-	-	-
-19			-	-	-	-		3 (5.9)	-	-
-18			-	-	-	-		-	-	10 (19.6)
-17	1 (2.0)		1 (2.0)	-	-	-		1 (2.0)	-	-
-16	-	2 (3.9)	-	-	-	4 (7.8)		-	-	-
-15	1 (2.0)	-	1 (2.0)	-	-	-		5 (9.8)	-	-
-14	-	-	-	-	-	-	7 (13.7)	-	-	-
-13	-	4 (7.8)	4 (7.8)	-	-	-	-	-	-	-
-12	-	-	-	-	-	-	-	-	-	-
-11	9.8	-	-	-	-	-	-	-	-	-
-10	1 (2.0)	2 (3.9)	1 (2.0)	-	-	-	-	-	-	-
-9	-	-	-	-	-	-	1 (2.0)	-	-	-
-8	1 (2.0)	-	1 (2.0)	1 (2.0)	-	1 (2.0)	-	-	-	-
-7	2 (3.9)	-	-	-	-	-	-	1 (2.0)	-	-
-6	-	4 (7.8)	3 (5.9)	-	-	1 (2.0)	1 (2.0)	-	-	-
-5	-	1 (2.0)	1 (2.0)	-	-	-	1 (2.0)	-	2 (7.4)	-
-4	4 (7.8)	2 (3.9)	1 (2.0)	-	-	2 (3.9)	3 (5.9)	1 (2.0)	-	-
-3	-	2 (3.9)	-	-	-	-	-	-	-	-
-2	4 (7.8)	2 (3.9)	3 (5.9)	-	-	-	1 (2.0)	-	1 (3.7)	-
-1	-	-	-	-	-	-	-	-	-	-
0	17 (33.3)	21 (41.2)	20 (39.2)	41 (80.4)	47 (92.2)	36 (70.6)	29 (56.9)	28 (54.9)	22 (81.5)	38 (74.5)
1	2 (3.9)	-	-	3 (5.9)		-	-	-		
2	3 (5.9)	3 (5.9)	3 (5.9)	2 (3.9)		2 (3.9)	-	-		
3	1 (2.0)	1 (2.0)	-	-		-	1 (2.0)	1 (2.0)		
4	2 (3.9)	-	4 (7.8)	-		1 (2.0)	-	2 (3.9)		
5	1 (2.0)	1 (2.0)	-	-		-	-	-		
6	-	-	1 (2.0)	-		1 (2.0)	-	-		
7	-	2 (3.9)	-	-		-	-	2 (3.9)		
8	-	-	-	-		-	-	-		
9	-	-	-	-		-	-	-		
10	-	-	1 (2.0)	-		-	-	-		
11	4 (7.8)	-	-	-		-	-	-		
12	1 (2.0)	-	-	-		-	-	-		
13	-	4 (7.8)	4 (7.8)	-		-	-	-		
14	-		-	-		-	5 (9.8)	-		
15	-		-	-		-	-	4 (7.8)		
16	-		-	-		-	-	-		
17	-		-	-		-	-	-		
18	1 (2.0)		-	-		-	1 (2.0)	-		
19			1 (2.0)	-		-	-	-		
20				-		1 (2.0)	-	-		
21				-			1 (2.0)	-		
22				2 (3.9)				1 (2.0)		
23				1 (2.0)						

Note. Δ = *T* score change from first to second ratings, OVR = Overactivity, UNR = Underactivity, ADH = Attention Deficit-Hyperactive, SA(P) = Solitary Aggressive (Provocative), SA(I) = Solitary Aggressive (Impulsive), OPD = Oppositional Defiant, DIF = Diffident, AVO = Avoidant, DEL = Delinquent, LEH = Lethargic (Hypoactive). SE_m = Standard Error of Measurement (McDermott, 1994, p. 46). Percents presented in parentheses.

Table 6
Stability of ASCA Syndromic Profile Based Classifications Using Generalized Distance Score (GDS) and $r_{p(k)}$ Methods and Discriminant Classifications

	P_o	P_c	κ	SE_κ	Z	p
GDS Method						
22 Syndrome Profiles	.41	.08	.36	.04	9.20	.00001
5 Broad Classifications	.61	.25	.48	.07	6.50	.00001
3 Broad Classifications	.84	.60	.61	.11	5.72	.00001
2 Broad Classifications	.94	.63	.84	.14	6.09	.00001
$r_{p(k)}$ *Method*						
22 Syndrome Profiles	.37	.08	.32	.04	8.20	.00001
5 Broad Classifications	.57	.25	.43	.07	5.86	.00001
3 Broad Classifications	.82	.59	.57	.10	5.55	.00001
2 Broad Classifications	.92	.62	.79	.14	5.70	.00001
Discriminant Classification						
Normal/SED (Linear)	.88	.73	.56	.13	4.19	.00003

Note. P_o = Observed Agreement, P_c = Chance Agreement.

For a copy of the *Excel* for the *Macintosh* spreadsheet templates for calculating the agreement over time (or between raters) for the 22 Syndromic Profiles; 5, 3, and 2 level broad classifications; and Discriminant Classifications; send a self addressed envelope, a *Macintosh* formatted 3.5 inch disk, and $5.00 to cover expenses.

References

Achenbach, T. M. (1991). *Manual for the Teacher's Report Form and 1991 Profile.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* Burlington, VT: University of Vermont.

Canivez, G. L. (1996). *Automated Syndromic Profile and Discriminant Classification Analyses for the Adjustment Scales for Children and Adolescents (ASCA) v2.0. Microsoft® Excel™* Spreadsheet Template for the *Apple® Macintosh™* Microcomputer. Charleston, IL: Author.

Canivez, G. L. (1998a). Automated syndromic profile and discriminant classification analyses for the Adjustment Scales for Children and Adolescents. *Behavior Research Methods, Instruments, & Computers, 30,* 732-734.

Canivez, G. L. (1998b). Automated calculation of nominal scale agreement (Kappa) statistics for ASCA syndromic profile classifications and discriminant classifications. *Microsoft Excel* Spreadsheet Templates for the *Apple Macintosh* Microcomputer. Charleston, IL: Author.

Cattell, R. B. (1949). r_p and other coefficients of profile similarity. *Psychometrika, 14,* 279-298.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Everitt, B. S., & Hay, D. F. (1992). *Talking about statistics: A psychologist's guide to data analysis.* New York: Wiley.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York: McGraw-Hill.

Kamphaus, R. W., & Frick, P. J. (1996). *Clinical assessment of child and adolescent personality and behavior.* Boston: Allyn and Bacon.

Kiess, H. O. (1996). *Statistical concepts for the behavioral sciences* (2nd ed.). Boston, MA: Allyn and Bacon.

Knoff, H. M. (1995). Best practices in personality assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-III* (pp. 849-864). Washington, DC: National Association of School Psychologists.

Landis, J. R., & Koch, G. C. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

McConaughy, S. H., & Ritter, D. R. (1995). Best practices in multidimensional assessment of emotional or behavioral disorders, *Best practices in school psychology-III* (pp. 865-877). Washington, DC: National Association of School Psychologists.

McDermott, P. A. (1988). Agreement among diagnosticians or observers: Its importance and determination. *Professional School Psychology, 3,* 225-240.

McDermott, P. A. (1994). *National profiles in youth psychopathology: Manual of Adjustment Scales for Children and Adolescents.* Philadelphia: Edumetric and Clinical Science.

McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment Scales for Children and Adolescents.* Philadelphia: Edumetric and Clinical Science.

McDermott, P. A., Watkins, M. W., Sichel, A. F., Weber, E. M., Keenan, J. T., Holland, A. M., & Leigh, N. M. (1995). The accuracy of new national scales for detecting emotional disturbance in children and adolescents. *The Journal of Special Education, 29,* 337-354.

McDermott, P. A., & Weiss, R. V. (1993). *Syndromic profile types for the Adjustment Scales for Children and Adolescents.* Philadelphia: University of Pennsylvania.

McDermott, P. A., & Weiss, R. V. (1995). A normative typology of healthy, subclinical, and clinical behavior styles among American children and adolescents. *Psychological Assessment, 7,* 162-170.Merrell, K. W. (1994a). *Assessment of behavioral, social, & emotional problems.* NY: Longman.

Merrell, K. W. (1994b). *Preschool and Kindergarten Behavior Scales*. Brandon, VT: Clinical Psychology Publishing Company.

Naglieri, J., LeBuffe, P. A., & Pfeiffer, S. I. (1993). *Devereux Behavior Rating Scale-School Form*. San Antonio, TX: The Psychological Corporation.

Reschly, D. J., & Ysseldyke, J. E. (1995). School psychology paradigm shift. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-III* (pp. 17-32). Washington, DC: National Association of School Psychologists.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service.

Salvia, J., & Ysseldyke, J. E. (1995). *Assessment*. Boston: Houghton Mifflin.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12*, 331-350.

Tatsuoka, M. M. (1974). *Selected topics in advanced statistics: An elementary approach: Number 3: Classification procedures: Profile similarity*. Champaign, IL: Institute for Personality and Ability Testing.

Trites, R. L., Blouin, A. G. A., & Laprade, K. (1982). Analysis of the Conners Teacher Rating Scale based on a large normative sample. *Journal of Consulting and Clinical Psychology, 50*, 615-623.

Watkins, M. W. (1997). *ASCAcalc* [Computer Software]. State College, PA: Author.

Watkins, M. W., & Canivez, G. L. (1997). Interrater agreement of the Adjustment Scales for Children and Adolescents. *Diagnostique, 22*, 205-213.